DOCUMENT RESUME

ED 280 310 FL 016 575

AUTHOR Wing, Barbara H.; Mayewski, Sandi F.
TITLE Oral Proficiency Testing in College-Level Foreign Language Programs. A Handbook for Foreign Language Departments.
INSTITUTION American Council on the Teaching of Foreign Languages, Hastings-on-Hudson, N.Y.
SPONS AGENCY Department of Education, Washington, DC.
PUB DATE Sep 84
GRANT G008201400
NOTE 36p.
PUB TYPE Information Analyses (070) -- Guides - Non-Classroom Use (055)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Classroom Techniques; *College Second Language Programs; Communication Skills; Curriculum Development; Departments; Evaluation Methods; Feasibility Studies; Higher Education; *Language Proficiency; Language Skills; *Language Tests; *Rating Scales; Second Language Instruction; Teaching Guides; *Test Use; Verbal Tests
IDENTIFIERS American Council on the Teaching of Foreign Langs; Educational Testing Service; Interagency Language Roundtable; *Oral Proficiency Testing

ABSTRACT

A handbook designed to acquaint college and university language departments with recent developments in oral proficiency testing gives an overview of the rationale, development, procedures, and implications of oral proficiency evaluation. The discussion is placed in the context of college and university language teaching. An overview section defines oral proficiency testing, looks at its importance in the academic setting, and outlines recent developments. A second section describes the oral proficiency evaluation process as outlined by the American Council on the Teaching of Foreign Languages (ACTFL) and the Educational Testing Service (ETS). The third section describes the academic uses of the ACTFL/ETS process, and the fourth section discusses the process' implications for curriculum development, specific courses, teaching techniques, and testing the other skills. The handbook also reports the results of a study concerning the feasibility of incorporating oral proficiency testing in college-level foreign language courses. The oral proficiency scales developed by the Interagency Language Roundtable and by ACTFL/ETS are appended. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

C. Linder

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

FL020803

ORAL PROFICIENCY TESTING IN COLLEGE-LEVEL FOREIGN LANGUAGE PROGRAMS

A Handbook for Foreign Language Departments

by

Barbara H. Wing and Sandi F. Mayewski
University of New Hampshire
Durham, New Hampshire
Directors, PROJECT OPT

PROJECT OPT
was funded by a grant (#G008201400)
from the International Research and Studies Program
of the U.S. Department of Education

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL016575

ACTFL Materials Center
579 Broadway
Hastings-on-Hudson, NY 10706

BEST COPY AVAILABLE

ACKNOWLEDGEMENTS

A handbook such as this one is not solely the work of its authors. It is a compilation of the ideas and experience of many people who have contributed to the development of oral proficiency testing over the past 30 years. Some of these people, whose work is available in print, are cited in the Reference section. Others, such as those who participated in the PROJECT OPT Study and other projects, are not identified for reasons of professional ethics. Still others, who are part of the expanding network of oral proficiency evaluation advocates, are found on campuses all around the United States. Each of these groups has and will continue to play a vital role in the process of defining and refining proficiency criteria.

Funds for this project were provided by a grant (G008201400) from the U.S. Department of Education, Title VI, International Research and Studies Program. The ideas and opinions expressed in this handbook are those of the authors and do not necessarily represent those of the funding agency.

Special thanks are due to Judith E. Liskin-Gasparro, Educational Testing Service, and F. William Forbes, University of New Hampshire, for their helpful editorial suggestions; to Betty LeCompagnon, University of New Hampshire, for the preparation of the final manuscript; and to C. Edward Scebold, Executive Director of ACTFL, for publishing and distributing this handbook.

Barbara H. Wing
Sandi F. Mayewski

September 1984

PREFACE

This handbook is designed to acquaint foreign language departments in colleges and universities in the United States with recent developments in oral proficiency testing in foreign language programs. It is intended to provide an overview of the rationale, development, procedures, and implications of oral proficiency evaluation in the context of the teaching mission of college and university foreign language departments. In addition, it reports the findings of PROJECT OPT, a federally funded study to determine the feasibility of incorporating oral proficiency testing in college-level foreign language courses.

It is important to note that this handbook is not intended to be a training manual but rather an introduction to the principles and techniques of oral proficiency testing. A Reference section provides additional information regarding the growing bibliography of printed sources and the expanding network of college and university professors with experience in oral proficiency evaluation. Chairpersons and professors who would like further information are advised to contact David V. Hiple, Project Director, American Council on the Teaching of Foreign Languages, 579 Broadway, Hastings-on-Hudson, NY 10706; Judith E. Liskin-Gasparro, Senior Examiner, Educational Testing Service, Princeton, NJ 08541; or the authors of this handbook at the University of New Hampshire, Durham, NH 03824. Training for faculty members can be arranged through ACTFL on the home campus or is available at periodic workshops held around the country.

TABLE OF CONTENTS

I. ORAL PROFICIENCY TESTING: AN OVERVIEW

1. What is ORAL PROFICIENCY TESTING?
2. Why is ORAL PROFICIENCY TESTING Important in Academic Settings?
3. What is Currently Happening in ORAL PROFICIENCY TESTING?

II. THE ACTFL/ETS ORAL PROFICIENCY EVALUATION PROCESS

1. What is the Process?
2. What is the ACTFL/ETS RATING SCALE?
3. How is the Interview Conducted?
4. How Can One Learn to Use this Process?

III. USES OF THE ACTFL/ETS ORAL PROFICIENCY EVALUATION PROCESS

1. How is the Process Being Used in Academic Settings?
2. What Wider Applications Relate to the Teaching Mission of Foreign Language Departments?

IV. IMPLICATIONS OF THE ORAL PROFICIENCY EVALUATION PROCESS

1. What are the Implications for Curriculum Development?
2. What are the Implications for Specific Courses?
3. What are the Implications for Teaching Techniques?
4. What are the Implications for Testing the Other Skills?

V. PROJECT OPT

1. What is PROJECT OPT?
2. What Answers does it Provide to Questions about ORAL PROFICIENCY TESTING in College-Level Foreign Language Courses?

VI. REFERENCES
APPENDICES

1. The Interagency Language Roundtable Oral Proficiency Scale
2. The ACTFL/ETS Oral Proficiency Scale

I. ORAL PROFICIENCY TESTING: AN OVERVIEW

WHAT IS ORAL PROFICIENCY TESTING?

Oral proficiency testing is an evaulation process designed to determine how well a person can *speak* a foreign language. The process consists of procedures that assess the ability of the speaker to use language effectively in communicative situations. Originally developed by the Foreign Service Institute of the U.S. Department of State in the 1950s, it has been adapted to academic settings in recent years in answer to a need expressed by many foreign language professors.

WHY IS ORAL PROFICIENCY TESTING IMPORTANT IN ACADEMIC SETTINGS?

Oral proficiency testing in academic settings is a recent development in foreign language teaching. It addresses a longstanding concern of college and university foreign language departments: the evaluation of oral skills. This concern is directly related to the teaching mission of foreign language faculties. The mission recognizes that the primary function of the department is to instruct majors in the foreign languages and literatures and that the secondary function is to provide language instruction for non-majors.

Achieving departmental goals related to the teaching mission requires coordination of curriculum, instruction, and evaluation. As recent foreign language texts attest, course content and teaching techniques emphasize the development of oral skills for communicative functions. Evaluation of these oral skills, however, has not been systematically planned and implemented. In contrast, techniques for evaluating grammatical knowledge, vocabulary acquisition, and reading comprehension are well-established components of the foreign language curriculum. These techniques are primarily pencil and paper tests that can be administered to groups of students and scored objectively. They provide valuable data relative to mastery of the morphological, syntactical, and lexical elements of the foreign language as well as the ability to comprehend and produce the written language. They do not, however, measure how well a student can communicate orally in the foreign language.

The ability to communicate orally in the foreign language is a primary goal of both students and faculties in foreign language departments. Implicit in this goal is the recognition that mastery of a foreign language benefits the individual and society in both humanistic and pragmatic ways. Thus, in the past two decades, many foreign language departments have expanded their focus to include courses and concentrations

that complement the traditional literary orientation. Likewise, students have recognized the ancillary value of language proficiency while preparing for careers in a broad range of other fields. The ability to communicate in a foreign language, even at limited proficiency levels, is increasingly recognized as a valuable asset in such areas as business and commerce, the service professions, travel and recreation, and the government.

Professors who are aware of the important role of oral proficiency have realized that this area must be evaluated with reliable and valid testing techniques. Such techniques provide a common standard that can be applied to all students, regardless of institution, level of study, or type of curriculum. Futhermore, they describe proficiency in terms that explain what students can actually say in the foreign language. This kind of analysis provides valuable information to the professor, the student, and prospective employers. Until recently, validated techniques were not available for wider use in foreign language departments in schools, colleges, and universities, although some individual professors had developed their own procedures.

WHAT IS HAPPENING IN ORAL PROFICIENCY TESTING?

In 1979, President Carter's Commission on Foreign Language and International Studies proposed, as part of a series of recommendations to improve foreign language instruction in the United States, that language proficiency goals and guidelines for evaluation be established. The suggestion was a timely one for two reasons.

In the first place, those foreign language professors who had themselves learned a foreign language in government-sponsored agencies, such as the Defense Language Institute, the Foreign Service Institute, or the Peace Corps, were aware of specific procedures to rate oral proficiency. These procedures, which have been refined over a period of 30 years, are now used regularly by the approximately 30 government agencies that constitute the Interagency Language Roundtable (ILR). They are the primary means to determine the functional language ability of candidates for the thousands of government positions that require a knowledge of a foreign language.

In the second place, the Educational Testing Service (ETS), a private educational agency that develops evaluation procedures and instruments, and the American Council on the Teaching of Foreign Languages (ACTFL), a professional organization of foreign language teachers, were already addressing the problem. In 1980, ETS, with funding from the U.S. Department of Education, undertook a project

entitled "A Common Metric for Language Proficiency." The result was an adaptation of the Interagency Language Roundtable proficiency rating scale to reflect the needs and realities of the academic setting. Subsequently ACTFL, in projects also funded by the U.S. Department of Education, sponsored workshops to train foreign language teachers in the use of the new scale. It also produced a series of proficiency guidelines, written by teams of college and university language professors, in the five areas of listening, speaking, reading, writing, and culture.

The response to the President's Commission has been a collaborative effort by ETS, ACTFL, and foreign language professors working in their own institutions to develop procedures, train testers, and investigate the impact and implications of utilization of oral proficiency tests. The most highly developed and commonly used procedure is the ACTFL/ETS Oral Proficiency Interview and Rating Scale, which will be discussed in Section II of this handbook.

Since 1982, the training of testers has taken place in one-day familiarization workshops and four-day intensive workshops in several parts of the country. Investigation of the impact and implications is a continuing process that has ramifications for the development of course and program goals, the design of instructional activities, and the development of texts and other teaching materials. Centers for research and training have been established, projects funded, and networks formed to provide a sound basis for continued development and implementation.

The successful development of techniques for oral proficiency testing constitutes a major step toward realizing the goal of communicative competence in foreign language learning and teaching. There are, however, more far-reaching implications than simply the evaluation of speaking skill. Research and development have been expanded to include the other modalities of listening, reading, writing, and culture. Consequently, proficiency as an organizing principle represents a vital and powerful concept for fulfilling the teaching mission of foreign language departments in the colleges and universities.

II. THE ACTFL/ETS ORAL PROFICIENCY EVALUATION PROCESS

WHAT IS THE PROCESS?

The ACTFL/ETS oral proficiency evaluation process is designed to provide a measure of *functional* language ability that describes how well a person can use the foreign language in real-life settings, regardless of how it has been learned or acquired. It is a modification, for academic settings, of the Interagency Language Roundtable (ILR) interview process. The process consists of a conversational interview and the rating of the resulting speech sample. The interview, which provides the language sample to be evaluated, is a highly structured conversation based on levels determined by the rating scale. Since an understanding of the rating scale is fundamental to conducting a successful interview, the scale will be discussed first.

WHAT IS THE ACTFL/ETS RATING SCALE?

The ACTFL/ETS rating scale is a descriptive instrument that defines levels of real-life language proficiency. It is a modification of the ILR scale, which was initially developed by the Foreign Service Institute (FSI). Figure I presents a comparison of the two scales.

The central cone of the figure represents the speaker's functional language ability, as demonstrated in the oral interview. The ILR scale, which has five levels of proficiency rating from 0 ("no functional language ability") to 5 (proficiency equivalent to that of an "educated native speaker") is indicated on the left side of the figure. The ACTFL/ETS scale, which is an expansion of the lower ends of the ILR scale, is indicated on the right side of the figure. Levels are identified by descriptors such as "S-1 (Speaking)" in the ILR scale which corresponds to "Intermediate" in the ACTFL/ETS scale. Each level and sub-division within levels (Low, Mid, High) is defined by a descriptive paragraph. These descriptive paragraphs comprise the rating scale. The complete ILR and ACTFL/ETS rating scales are found in the Appendices.

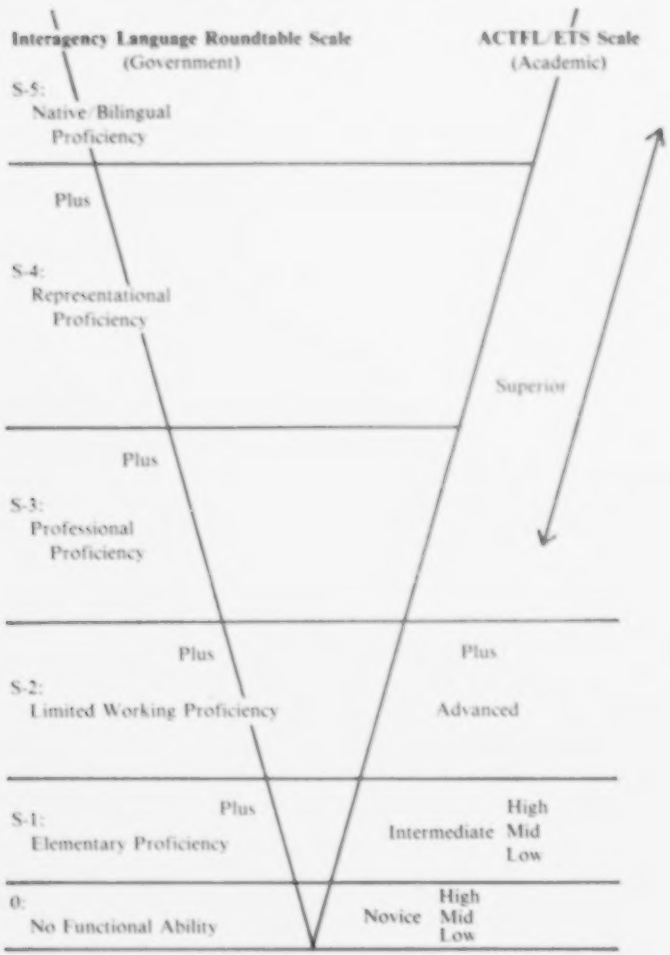
Based on research and experience with the ILR scale, ACTFL and ETS expanded the lower end of that scale (0, "no functional ability," to 2, "limited working proficiency") to achieve a more accurate measure of discrimination in the ranges most commonly achieved by students in classroom learning situations. Thus, the ACTFL/ETS rating scale consists of ranges of increasingly larger increments on a scale of Novice-

Low to Superior (0 to 3 on the ILR scale). Each level and its related subdivisions are characterized by a single-sentence identifier that is followed by a descriptive paragraph. For example, a "Novice-Low" is

> *Unable to function in the spoken language.* Oral production is limited to occasional isolated words. Essentially no communicative ability.

Figure 1

LANGUAGE PROFICIENCY SCALES



An "Advanced Plus," on the other hand, is

> *Able to satisfy most work requirements and show ability to communicate on concrete topics relating to particular interests and special fields of competence.* Often shows remarkable fluency and ease of speech, but under tension or pressure language may break down. Weaknesses or unevenness in one of the foregoing or in pronunciation result in occasional miscommunication. Areas of weakness range from simple constructions such as plurals, articles, prepositions, and negatives to more complex structures such as tense usage, passive constructions, word order, and relative clauses. Normally controls general vocabulary with some groping for everyday vocabulary still evident.

At each level of the scale, function, context, and accuracy requirements comprise the description of proficiency. *Function* refers to the tasks accomplished, attitudes expressed, and tone conveyed by the student. *Context* refers to the topics, subject areas, activities, and jobs the student is able to address. *Accuracy* refers to the acceptability, quality, and accuracy of the message conveyed by the student. Each interview is rated by taking into account all three of these areas. The Functional Trisection (see Figure II) provides a succinct summary of the levels of the scale.

A rating is assigned by determining the highest level of sustained performance as evidenced by patterns of strengths and weaknesses. It should be noted that two different speakers may vary in the way they demonstrate content, function, and accuracy, but still fall in the same level of proficiency. Furthermore, the levels of the scale represent ranges of increasing magnitude rather than equidistant points on a fixed line. Thus, as one progresses up the scale, the quantity and quality of language increase geometrically, not arithmetically. This progression is graphically illustrated in the cone-shape of Figure I, where each level is represented by increasingly larger segments.

Rating is normally done from a tape after the interview has been conducted. This removes the tester from the pressure of simultaneous interviewing and rating. It also allows for review of the interview as many times as necessary and with some psychological distance. Furthermore, the rating can be verified by other testers.

Figure II

FUNCTIONAL TRISECTION OF ORAL PROFICIENCY LEVELS

Oral Proficiency Level	Function (Tasks accomplished, attitudes expressed, tone conveyed	Context (Topics, subject areas, activities and jobs addressed	Accuracy (Acceptability, quality and accuracy of message conveyed)
5 (Superior)	Functions equivalent to an educated native speaker (ENS).	All subjects.	Performance equivalent to an ENS.
4 (Superior)	Able to tailor language to fit audience, counsel, persuade, negotiate, represent a point of view and interpret for dignitaries.	All topics normally pertinent to professional needs.	Nearly equivalent to an ENS. Speech is extensive, precise, appropriate to every occasion with only occasional errors.
3 (Superior)	Can converse in formal and informal situations, resolve problem situations, deal with unfamiliar topics, provide explanations, describe in detail, offer supported opinions, and hypothesize.	Practical, social professional and abstract topics, particular interests, and special fields of competence.	Errors never interfere with understanding and rarely disturb the native speaker. Only sporadic errors in basic structure.
2 (Advanced)	Able to fully participate in casual conversations, can express facts, give instructions, describe, report, and provide narration about current, past and future activities.	Concrete topics such as own background, family, interests, work, travel, and current events.	Understandable to native speaker *not* used to dealing with foreigners sometimes miscommunicates.
1 (Intermediate)	Can create with the language, ask and answer questions, participate in short conversations.	Everyday survival topics and courtesy requirements.	Intelligible to native speaker used to dealing with foreigners.
0*	No functional ability.	None.	Unintelligible.

*Novice Level is not discussed here.

Used with permission of Pardee Lowe, Jr.

8

HOW IS THE INTERVIEW CONDUCTED?

The interview is a structured conversation that enables the tester to determine what the student can do and what the student cannot do when speaking the target language. Since the object of the interview is the assessment of language proficiency, a ratable sample must be obtained from the speaker. It follows, therefore, that each question posed by the tester must have a purpose. Structured questioning not only shortens the length of the interview, but also improves the sample, thus allowing the tester to evaluate more accurately.

To ensure that questioning is structured, the interview is divided into four phases. These four phases are: warm-up, level check, probes, and wind-down. Such a structure avoids purposeless questioning that lengthens the interview and generally results in an unratable sample.

The *warm-up* is designed to put the speaker at ease and, if necessary, to reacquaint him/her with the target language. It consists of social amenities and very simple conversation, including salutations and introductions. The warm-up is concluded when the interviewer has obtained a preliminary indication of the student's level and when the student seems relaxed and comfortable with the testing situation.

The *level check* provides information regarding what the speaker *can* do. It is designed to find the highest level at which the student can sustain a speaking performance. Its purpose is to verify the estimate of level obtained at the end of the warm-up. During this stage the tester checks a number of topics for content and vocabulary, listens for fluency, and assesses pronunciation and grammatical accuracy. When the speaker is able to perform the functions of the estimated level with suitable accuracy and content, the solid floor of successful performance has been established. This floor represents what the speaker *can* do consistently.

Probes provide information regarding what the speaker *cannot* do. In contrast to "level checks," which reveal what a speaker can do consistently, thus determining the floor of the performance, "probes" are used to find the ceiling for the rating. The probe stage takes the student above the sustained level several times in different ways, exploring a variety of contexts as well as a variety of functions. The objective is to test the speaker to the highest level of proficiency. When the speaker can no longer communicate effectively, this objective has been realized and the testing part of the interview is complete.

The *wind-down* is designed to return the speaker to the level at which he/she functions most accurately. This is done to provide the speaker

with a feeling of accomplishment at the close of the interview. The tester often returns briefly to a topic discussed previously and on which the speaker had been able to converse with a considerable degree of confidence and ease. The interview generally ends, as it began, with the appropriate social amenities.

During the course of the interview, testers are faced with the task of generating questions appropriate to the various stages of the interview and, concurrently, appropriate to the student's apparent level of language proficiency. They often prepare *types* of questions suitable to each level before the interview. The actual questions, however, are developed as the interview progresses, according to the information that the tester is gathering about the speaker.

In addition to following a question/answer pattern, speakers are asked to role-play situations with the interviewer for levels above Novice. These role-plays are designed to give the speaker an opportunity to demonstrate how well he/she can get into, through, and out of a situation that one might normally encounter when using the foreign language. They require that the speaker ask questions, sustain a conversation and, at higher levels, deal with complications.

HOW CAN ONE LEARN TO GIVE THIS TEST?

The interviewer is not a teacher but rather an evaluator. The two roles must be kept separate when conducting interviews. Although many foreign language professors have had training in pedagogy and may possess a native, or near-native, command of the language they teach, it does not follow that they can automatically conduct an oral proficiency interview. Careful training and a substantial number of practice interviews are necessary before even the best teacher can elicit a ratable sample from a speaker and subsequently rate the sample accurately. Such training is available in workshops conducted by individuals trained by ACTFL and ETS.

An introduction to the principles and procedures of the ACTFL/ETS oral proficiency interview and rating scale can be obtained by attending a familiarization workshop of one or two days duration. Participation in an intensive workshop of four or five days duration, followed by post-workshop training (in which participants conduct practice interviews that are critiqued by their trainers), is a highly recommended precondition for certification as an interviewer and rater. For information regarding workshops and tester certification, consult the sources mentioned in the Preface of this handbook.

III. USES OF THE ACTFL/ETS ORAL PROFICIENCY EVALUATION PROCESS

The ACTFL/ETS oral proficiency interview is a modification of the Interagency Language Roundtable interview process that reflects reasonable expectations for proficiency development in academic settings. By expanding the lower ends of the ILR scale (0-3), ACTFL and ETS have provided a more accurate measure of discrimination in the ranges most commonly achieved by students in classroom learning situations. Since the ILR and ACTFL/ETS processes and rating scales are mutually compatible, they are appropriate for use in a variety of settings and for a variety of functions.

HOW IS THE PROCESS BEING USED IN ACADEMIC SITUATIONS?

The techniques and principles underlying the oral proficiency interview serve a broad range of interests for those involved in language instruction in colleges and universities. Some of the most common applications, as reported in the March, 1984 issue of *POPT*, the newsletter of PROJECT OPT, are:

1. *Placement in foreign language courses:* Appropriate placement of incoming students in the elementary, intermediate, and advanced level courses is necessary for both the individual and for the class in order to provide optimal learning conditions. Many institutions administer a written placement examination to determine if students have the necessary command of vocabulary and structure to enroll in specific courses. In cases where oral proficiency is an important factor, such as conversation courses, or where the score on the written test falls in the overlapping range of more than one possible course, the process is useful to suggest which course will be most beneficial. Currently it is being used for this purpose at Bates College (Maine), Norwich University (Vermont), and Cabrillo College (California).

2. *Evaluation of oral skills in specific foreign language courses:* Some professors use the process at the beginning of the semester in intermediate and advanced courses to determine entry-level proficiency of their students. There are two benefits to such utilization. The interview provides an objective measure of what students can do prior to receiving further instruction. The resultant evaluation can be used for diagnostic purposes since the professor has obtained a valid assessment of strengths and weaknesses. The preliminary interview can also be compared with one conducted at the end of the course. It is important to note, however, that the levels of the ACTFL/ETS scale *do not* parallel

course levels and thus a change in proficiency level may not be a reasonable expectation. Furthermore, the inteview provides the opportunity for professor and student to become acquainted in a less formal environment than in the classroom. This personal contact at the beginning of the course lays the groundwork for developing an in-class atmosphere that fosters real communicative use of the foreign language.

In addition to entry-level assessment, the oral interview is also used as *part* of the criteria for determining a final grade in specific courses. It is important, however, to recognize that many factors, including achievement-based measures, determine grades. As stated above, the levels of the ACTFL/ETS scale do not parallel course levels. Neither do levels of proficiency equal course grades. Examples of this type of application can be found in a German advanced composition and conversation course at Amherst College (Massachusetts), in elementary and intermediate French courses at the University of Illinois (Urbana-Champaign), in elementary French courses at the University of Wisconsin (Madison), and in Russian courses at the University of Vermont.

3. *Selection of students for a study-abroad program and determination of the effects of such a program:* At Central College (Pella, IA), students are required to reach a strong 1+ or 2 on the ILR scale (Intermediate High or Advanced on the ACTFL/ETS scale) for acceptance into the JYA program in Granada, Spain. At Keene State College (New Hampshire), students of French, German, and Spanish are tested before and after the study abroad program. Russian students attending the University of New Hampshire Associated Academic Programs in Leningrad are tested, for placement purposes, at the beginning of the summer program and then again at the end to determine progress in proficiency development. Likewise, Russian students enrolled in programs sponsored by the American Council of Teachers of Russian at the Pushkin Institute in Moscow are tested and receive an oral proficiency rating both prior to and after attending the semester and year courses.

4. *Fulfillment of specific program requirements:* With the increasing demand for accountablility in education, it is important that foreign language departments be able to verify that students have completed such requirements as:

a. The foreign language requirement: In most institutions that have a foreign language requirement, fulfillment of that requirement is achieved by completing a specific course or series of courses. This approach leaves decisions regarding content, skill development, and level of achievement up to the department, or in some cases, to individual professors. It does not necessarily focus on what students should be

able to do in the language, once they have completed the necessary course credits, nor does it provide a means of comparing what students at different institutions can do. The inclusion of an oral proficiency interview in the requirement criteria, such as at the University of Pennsylvania, where students are tested in French, German, Arabic, and Russian, assures that the oral skills have been evaluated in a way that has validity and meaning beyond the individual course or institution.

b. Requirements for successful completion of the foreign language major: Similar to the use of the oral proficiency evaluation process for non-majors fulfilling the language requirement is its application in dealing with majors and minors. Oral proficiency is a critical factor both in successful completion of upper-division courses and in post-graduation career uses of the foreign language: therefore it must be evaluated in an objective manner. The interview provides a mechanism for assessing proficiency without impinging on the departmental prerogative for determining what students should study to complete the major. It also provides a standard that can be interpreted by prospective employers and graduate schools.

c. Requirements for teacher education programs: The U.S. educational system has been the focus of over two dozen important reports in the past three years. A recurring theme has been the need for more competent teachers in all disciplines. One significant reaction to the often critical remarks regarding teacher preparation is the mandating by state legislatures of competency examinations for teacher education students. In foreign languages, several states have developed written proficiency examinations or are requiring existing standardized national tests such as the National Teachers Examination. A pilot project in Texas, funded by the U.S. Department of Education and conducted under the auspices of the Texas Education Agency, the Texas Foreign Language Association, and ACTFL, is investigating what oral proficiency levels are appropriate for beginning teachers. By 1986 it is hoped that an oral proficiency requirement will be part of the foreign language teacher certification process. In Central College (Pella, IA), teacher education students must achieve a rating of 3 on the ILR scale, equivalent to a "Superior" on the ACTFL/ETS scale.

As the specific academic applications described above illustrate, the Oral Proficiency Interview is being used for several purposes in foreign language departments in colleges and universities. A familiarization with the ACTFL/ETS scale and proficiency guidelines is providing foreign language educators with a better sense of specific needs of students at the outset of each semester, with a more defined goal to guide their course design during the semester, and with a markedly more delineated

and standardized method of evaluation at important checkpoints in their programs.

WHAT WIDER APPLICATIONS RELATE TO THE TEACHING MISSION OF FOREIGN LANGUAGE DEPARTMENTS?

Because the ACTFL/ETS oral proficiency evaluation process is based on the same principles and procedures as the Interagency Language Roundtable interview and scale, the potential exists for extensive use outside of academic situations. Ratings that are determined by *certified* evaluators can become part of the student's transcript and/or dossier to indicate the level of proficiency achieved in the foreign language. These ratings, which are defined in terms of the content, function, and accuracy of the person's language skills, provide an objective and standardized appraisal of how well that person can communicate. They convey clearly to employers in industry, commerce, social services, travel and entertainment, government, and other career fields requiring a functional knowledge of a foreign language just what a student applying for a job can do.

It cannot be too strongly emphasized that such extra-academic uses of the oral proficiency evaluation process are only valid *when* the interview and rating is done by a *certified* tester. The process, while seemingly simple to comprehend, requires extensive training and practice in all phases of interviewing and rating. To preserve the integrity of the system and the ratings, individuals and institutions must resist the temptation to use ratings that have not been obtained by a certified evaluator on official documents that become part of a person's academic file. Procedures have recently been developed by ACTFL in conjunction with ETS and ILR, to grant official tester certification, which is renewable at specified intervals, upon completion of appropriate training in interviewing and rating.

IV. IMPLICATIONS OF THE ORAL PROFICIENCY EVALUATION PROCESS

Implicit in the implementation of a precise and standardized method of evaluation of oral skills is a wide-ranging effect on the future of foreign language programs in the United States. It is expected that once the oral proficiency interview is adopted on a broad scale by both secondary and higher education institutions across the country, it will have substantial impact on foreign language curricula as a whole, on specific course goals and objectives, and on the individual teaching and testing techniques of language professors. The implications of this impact are discussed below.

WHAT ARE THE IMPLICATIONS FOR CURRICULUM DEVELOPMENT?

Curriculum development concerns the definition of goals and objectives of the foreign language program. It focuses on what students will be expected to know (content) and do (skills) as a result of the instruction they receive in the language, literature, and culture and civilization courses that constitute the departmental offerings. Use of oral proficiency evaluation as an organizing principle for curriculum is likely to have the following implications:

1. The teaching mission of foreign language faculties can be reexamined and subsequently redefined in accord with a proficiency-oriented approach to language instruction. If courses are taught with a focus on building proficiency, then students can be expected to demonstrate functional use of the language in addition to completing course requirements.

2. Traditional level expectations, i.e. elementary, intermediate, advanced, can be redefined in accordance with empirically derived data, provided by the oral proficiency interview, relative to student ability to communicate orally. The definitions of the proficiency levels of the ACTFL/ETS scale can provide a reasonable structure of expectations for skill development. Thus, articulation between courses and levels can be smoother since expectations will be more clearly defined.

3. An attempt can be made to achieve more standardized proficiency expectations for graduating seniors and pre-service teacher candidates across the nation. Such standardization would ensure more uniform performance criteria in use of the foreign language but would leave specific content decisions to the individual departments.

WHAT ARE THE IMPLICATIONS FOR SPECIFIC COURSES?

The curriculum of the foreign language department is realized and implemented through the specific courses that make up the program. The oral proficiency evaluation process is likely to have the following implications within specific courses:

1. Courses within each level, i.e. elementary, intermediate, and advanced, can be redesigned, using the level descriptions of the ACTFL Proficiency Guidelines (the rating scale), to include course-specific mastery levels. Such an application has already been implemented in an advanced French conversation course at Northwestern University in which course goals included demonstrating the ability to "narrate and describe in the past, present, and future." These goals correspond to the Advanced level of the ACTFL/ETS scale.

2. Courses designed specifically as a means of augmenting the conversational skills can play a more prominent role in the curriculum. In addition, such courses can be redesigned, based on a common standard of proficiency, with the aim of defining more realistic expectations with regard to what a student can actually say in the foreign language.

3. Oral achievement tests, patterned on the oral proficiency interview and based on course-specific material, can become a valuable tool in evaluating student ability to communicate. By administering more oral tests in their courses, professors will convey an important message to students. That message is that the department and the professor take seriously their responsibility to help students develop real proficiency in the foreign language.

WHAT ARE THE IMPLICATIONS FOR TEACHING TECHNIQUES?

During the past 40 years, techniques for teaching foreign languages have increasingly emphasized the oral and communicative aspects of classroom learning. As a result of research findings that demonstrate a positive correlation between meaningful activities and amount of language learning, foreign language teachers have realized the importance of creating a classroom environment in which students speak to each other about topics that are significant to the individual and to the group. The oral proficiency interview, which is itself a structured conversation in which the speaker is led to increasingly higher levels of linguistic and discourse function, has the following implications for instructional techniques:

1. The use of the target language at all levels of instruction can play a more prominent role in the instructional process. Professors whose goals include teaching for proficiency realize the importance of providing meaningful input in the target language. They communicate with their students in the target language, at a level that challenges but does not frustrate, on topics appropriate to the learning tasks. They also are aware of the fact that, given the appropriate input, skill in listening comprehension develops more rapidly than skill in speaking and therefore do not insist on extremely high levels of accuracy when students are attempting to express their ideas and opinions.

2. To promote meaningful discourse among students, activities in pairs or small groups can replace the more typical student-teacher exchanges that have characterized more traditional language-learning approaches. In such activities, students assume responsibility for initiating

and maintaining conversation in addition to the typical pattern of responding to direct questions usually posed by the professor.

3. "Situations," such as those used in a structured oral proficiency interview, can be an effective method for developing a conversation similar to one the students might experience in the foreign country. Since these role-play situations reflect increasingly more complicated instances of typical daily encounters, they require use of the language at increasingly complex levels of function, content, and accuracy. In addition to gaining practice in meeting the linguistic challenges of such encounters, students experience a sense of satisfaction and accomplishment when they "get into, through, and out of" a simple survival situation or one with a complication. These feelings help to build confidence and motivation to continue the process of learning the foreign language.

WHAT ARE THE IMPLICATIONS FOR TESTING THE OTHER SKILLS?

The oral proficiency interview and rating scale were designed initially to evaluate speaking primarily and, by virtue of its utility in the process, listening comprehension secondarily. The development of the *ACTFL Proficiency Guidelines* in the skills of listening, reading, and writing has already generated interest and activity in the evaluation of proficiency in those areas. The Symposium on Receptive Language Skills, conducted by ACTFL in November, 1983, and reported in the September, 1984 issue of *Foreign Language Annals*, is evidence of the relevance of the concept of proficiency-based evaluation for all the skills. Some of the important implications include:

1. Testing in the other skill areas, as well as in speaking, will be focused on functional language use in order to determine what students can understand, read, and write in circumstances that reflect real language for communication use. Such testing is more global than specific in nature and is designed to yield ratings that correspond to those in the oral proficiency rating scale.

2. Such testing will provide further evidence that students progress at different rates in the acquisition of different skills. For example, a typical total proficiency profile might indicate that a student is an "Intermediate- High" in listening and reading comprehension, but an "Intermediate-Mid" in writing and possibly an "Intermediate-Low" or even a "Novice-High" in speaking. These ratings should not surprise professors since they reflect, to some extent, a natural order of

language acquisition and, perhaps to a greater extent, typical priorities in classroom language learning. Professors who want to improve their students' ratings in certain skill areas such as speaking, in particular, can subsequently make changes in the materials and instructional techniques they use to help students achieve the desired goals.

V. PROJECT OPT

WHAT IS PROJECT OPT?

PROJECT OPT is a research and training project on oral proficiency testing in college-level foreign language courses conducted at the University of New Hampshire under a grant from the U.S. Department of Education in 1982-84. It was designed to achieve three objectives:

1. to determine the feasibility of incorporating specific techniques of oral proficiency evaluation in foreign language courses at the college level;

2. to train professors to use the ACTFL/ETS oral proficiency interview procedures;

3. to create a network in northern New England colleges and universities of professors interested in oral proficiency evaluation.

The project began with an intensive four-day training workshop in which 24 professors of French, German, Russian, and Spanish were trained to use the ACTFL/ETS oral proficiency interview and rating scale. Upon returning to their home campuses, the workshop participants conducted 24 interviews, a sample of which was critiqued by experienced evaluators from such institutions as the Educational Testing Service, the Defense Language Institute, and Georgetown University. At a follow-up workshop six months later, participants met with the evaluators and addressed specific problem areas. Throughout the training period, data, which were subsequently used in the analysis reported below, were collected from the participants.

Nine research questions associated with the principal objective of determining the feasibility of incorporating specific techniques of oral proficiency evaluation in foreign language courses at the college level were investigated. These questions are representative of what faculty members often ask about the procedures and results of oral proficiency evaluation. A brief summary of the answers to these questions follows.

WHAT ANSWERS DID PROJECT OPT PROVIDE TO QUESTIONS ABOUT ORAL PROFICIENCY TESTING IN COLLEGE—LEVEL FOREIGN LANGUAGE COURSES?

*1. Can professors learn to conduct and rate oral proficiency interviews effectively after completing an intensive four-day workshop and a supervised practicum that includes a critique of 24 interviews?

The answer to this question is a qualified yes. When the study was conducted, a procedure for certifying new testers had not yet been defined. Thus it was not possible, at that time, to state definitively if the type and length of training was adequate to lead to certification of a high percentage of the workshop participants. However, the evaluators who critiqued the tapes, did judge several criteria related to the participants' performance and potential. They gave *high* ratings to 57% for interviewing techniques, 30% for rating techniques, 70% for potential as proficient interviewers, and 65% as proficiency raters. A uniform 13% received *low* ratings in all categories, with the remaining percentages falling in the *average* category.

Two conclusions can be drawn from the data. In the first place, it is apparent that the intensive training and the interviewing practice of at least 24 interviews are *minimal* requirements for tester certification. In fact, most experienced testers agree that many more interviews are necessary before one becomes proficient and comfortable. In the second place, it is obvious that rating is more problematic than interviewing. The distinctions within and across levels require a thorough understanding of the scale that only comes with experience in using it.

Since a tester certification procedure now exists and is being implemented by ACTFL/ETS, a more definitive answer to this question will be possible. Participants who complete the initial training and practicum now submit tapes and ratings to ACTFL for validation. As a result of this process, data can be gathered to determine the rate of success as well as the kind of supplementary training needed.

*2. What demands, in terms of time needed for preparation, administration, and scoring, does use of the oral proficiency interview make upon professors?

The amount of time required to conduct an oral proficiency interview, which is of necessity an individual process, is a matter of legitimate concern. The data revealed that beginning interviewers spent an average of 56 minutes arranging for, conducting, and then

reviewing the taped interview to rate it. Almost a quarter of this time was devoted to obtaining the specific type of interviewees needed for the project, with the remaining three-quarters spent in interviewing and rating.

As testers gain experience in choosing appropriate questioning techniques, the amount of time needed to obtain a ratable sample is reduced considerably. While there are no set time limits, a Novice-level interview generally lasts from 5-7 minutes, an Intermediate-level about 10-12 minutes, and an Advanced-level or Superior-level from 15-20 minutes at the most. Taping the interview provides a record that can be used for delayed rating or to check a rating determined immediately upon completion of the interview.

*3. What special conditions, facilities, and equipment do professors need to use the oral proficiency interview?

Participants did not indicate a need for special conditions, facilities, and equipment, i.e. those not normally available in the department, for conducting the interviews. Three-quarters of the participants used their own offices to conduct the interviews while the rest used the language laboratory. Portable tape recorders proved adequate though some professors reported problems with background noise and low-quality tapes. Students did not appear to be negatively affected by the presence of the tape recorder. It is important to make sure that there are no telephone calls or similar interruptions during the course of the interview.

*4. What impact does the oral proficiency interview have upon the curriculum, the instructional materials, and the procedures used to achieve the goals of the course?

The primary impact was perceived to be more emphasis on developing communicative skills by using proficiency-oriented texts, task-oriented instructional activities (especially role-play "situations") and authentic recorded materials in class. (For a more complete discussion of specific applications, see Section IV.)

*5. What presage and context variables are associated with successful use of the oral proficiency interview?

This question was designed to determine what relationships, if any, exist between (1) teacher characteristics, professional training, and experience, and (2) success in interviewing and rating. Because of the small sample size in each of the four languages represented in the study (French, 8; German, 4; Russian, 4; and Spanish, 9) the analysis was not carried out. Experience has shown, however, that

interviewers should themselves be at Level 3, and preferably 3+, on the ILR scale in order to obtain a ratable sample and to make accurate rating decisions along the full range of the ACTFL/ETS rating scale in academic situations.

*6. What problems do professors encounter when using the oral proficiency interview in a foreign language course?

The three most frequently mentioned problems concern (1) assuming the role of tester especially in the interview; (2) using probes appropriately in the interview; and (3) assigning an appropriate rating, especially in the Intermediate range.

(1) There are significant differences between the role of teacher and the role of tester. In the classroom, language teachers use several typical behaviors to encourage and reinforce learning. These include cueing responses, prompting students who hesitate, repeating student responses for correction or reinforcement purposes, and reacting with evaluative comments, especially when students respond correctly. In the interview, language testers must avoid these kinds of behaviors for a very specific reason: obtaining a ratable sample of the student's language proficiency depends upon an objective approach to interviewing that requires that the student alone show what he/she can and cannot do. Prompting, repeating, and encouraging by means of judgmental feedback on the part of the interviewer do not allow the student to do this. Thus, teacher-testers must learn to wait for answers without prompting, accept what students say without correcting or evaluating what is said, and permit students to reach the point where they can no longer express themselves comfortably. They must do these things, however, in a non-threatening manner that does not intimidate. Teacher-testers report that it is most difficult to accomplish this when one interviews one's own students (see 8 below).

(2) In the interview, probes are used to find out whether or not the speaker can go beyond the estimated level of proficiency. They serve to establish the "ceiling" or highest level at which the speaker can function consistently by taking the speaker to "linguistic breakdown," a point at which he/she can no longer convey meaning effectively. An interview should include at least three probes in order to show conclusively that the speaker cannot sustain speech at a higher level. Beginning testers often feel awkward putting speakers in such an uncomfortable position once, let alone three times during an interview. They also are not sure of the appropriateness

of their probes and may skip a level with the result that the speaker is not tested at a level which he/she might have been able to manage.

(3) Rating the interview requires considerable familiarity with the ACTFL/ETS rating scale and the distinctions between and within levels. First, it is extremely important that raters recognize the characteristics of each of the four base levels (Novice, Intermediate, Advanced, and Superior) in the system. Errors in rating decisions across these major borders indicate significant problems with either the sample or the matching of the sample to the definitions. Then, it is important that raters understand the different degrees of proficiency (Low, Mid, High) within the levels. While errors in assigning these ratings within levels are not as critical as those across levels, they are a matter of concern and should be corrected with additional practice. It should be noted, however, that the ACTFL/ETS definitions are still in a provisional, or developmental, stage, and some are acknowledged to need revision to eliminate possible ambiguities.

*7. What further training is necessary to help professors test oral proficiency in their courses?

Participants in the study agreed that the training sessions and the 24 critiqued interviews provided them with a basic understanding of the principles and procedures of the evaluation process. They acknowledged that increased confidence and accuracy in interviewing and rating are a function of experience. They recommended more observation of live or videotaped interviews and additional training, through shorter follow-up workshops, to determine plus points in levels. The evaluators who critiqued the participant tapes indicated a need to provide more training in estimating proficiency during the interview to make appropriate choices of interviewing techniques such as probes.

*8. What differences, if any, appear in the interviewing and rating techniques of professors when interviewing their own students and students whom they do not know?

A critical question is whether or not professors can obtain valid and reliable ratings when interviewing their own students. In many foreign language departments, the number of faculty members teaching a particular language may be limited to one or two professors and thus, if the students are to be interviewed and rated, it must be done by someone who is already familiar with what the students have done in class. The study provided data regarding the participants' perceptions of rating their own students and the evaluators' perceptions.

While acknowledging that interviewing one's own students is more difficult than interviewing those one does not know, over two-thirds of the participants in the study believed that interviews with their own students were as valid and reliable as those with unknown students. Twice as many felt more comfortable with their own students then with unknown students (3); however, a similar number noted that they had to resist slanting the interview toward known vocabulary and grammar. A third of the participants were surprised at the differences between the ratings and class performance with the usual expectation being a higher rating.

The evaluators who critiqued the participant tapes affirmed that, regardless of whether or not the professors were interviewing known or unknown students, 95% of the interviews constituted ratable samples. This finding indicates that, in the opinion of these experienced raters, it is possible to interview one's own students in a valid and a reliable manner by objectively following the guidelines.

When critiquing the participants' performance with regard to assigning a rating, the evaluators found more to question. Based on a stratified sample of all the interviews, the evaluators agreed with 39% of the participants' ratings and disagreed with 61%. When they disagreed, they raised the rating 47% of the time and lowered it 53% of the time. In 82% of the cases of disagreement, the magnitude was one step in the rating scale (from a Low to Mid, for example). These findings indicate that beginning raters are likely to rate high or low, by one step, but are generally close to the rating assigned by experienced raters.

When the data were analyzed to determine differences in rating known and unknown students, the evaluators agreed with the rating of known students 26% of the time and with unknown students 54% of the time. With more than twice the probability of agreement of rating unknown students between experienced and inexperienced raters, it appears that factors associated with previous classroom contact with students do affect a rating given by beginning testers. It is interesting to note, however, that when evaluators disagreed with ratings, they raised them 44% of the time for known students but only 38% of the time for unknown students. It would seem that professors might be compensating for their experience with known students by giving lower ratings.

The statistics support the participant consensus that it is more difficult for beginning testers to rate their own students than to rate students who have not studied with them. A reasonable solution to the dilemma is to have students interviewed by qualified testers, even

if the student has studied with the interviewer, and then have the tapes rated by independent evaluators. When a larger pool of certified testers is available, it will be interesting to continue the study of the validity and reliability of rating one's own students.

*9. What proficiency levels are associated with typical levels of instruction (beginning, intermediate, advanced) at the college level?

Participants in the study interviewed students enrolled in beginning, intermediate, and advanced language classes and in literature and culture classes taught in the foreign language. The resultant data provides some preliminary information on expected levels of proficiency in each of the typical levels of instruction. It should be noted that the sample (beginning students = 35, intermediate students = 71, advanced students = 28, literature and culture students = 20) is small and includes all four languages (French, German, Russian, and Spanish) of the project. In addition, the ratings were determined by beginning testers involved in a training project. As discussed above, experienced evaluators disagreed with these ratings 61% of the time and lowered the rating 53% of the time. Half of these disagreements involved major boundary decisions (from Intermediate to Advanced, for example) and half involved decisions within levels (from Mid to High, for example).

The proficiency levels in each level of instruction were:

1. Beginning language courses (1st and 2nd semester)
   - Spanish (n = 20): Novice-High
   - Intensive Russian (n = 15): Intermediate-Mid
2. Intermediate language courses (3rd and 4th semester)
   - French (n = 32): Intermediate-Mid
   - German (n = 23): Advanced
   - Russian (n = 15): Intermediate-Mid
3. Advanced language courses (5th and 6th semester)
   - French (n = 8): Intermediate-High
   - Spanish (n = 20): Intermediate—Mid
4. Literature and Culture courses
   - French (n = 8): Intermediate-High
   - Spanish (n = 12): Intermediate-High

While it may seem natural to equate level of course with level of proficiency, this tendency must be avoided, especially as the level of instruction advances. Tentative as the data is, it does reflect an important characteristic of the ACTFL/ETS scale: the ranges on the

scale are not fixed increments on an absolute scale but rather increasingly larger segments that require progressively greater command of function, content, and accuracy of language use. Thus, while the Beginning Spanish students were rated predictably in the Novice range, all the other students, with the exception of the Intermediate German students, were rated in the Intermediate range. The fact that 48 of these students were enrolled in what, for instructional purposes, are called "advanced" courses does not mean that they demonstrate proficiency at the Advanced levels.

As more testers are certified and more students are evaluated using the oral proficiency interview, statistics related to the relationship of course level and proficiency level will be more reliable. Such statistics will be useful to departments in colleges and universities in determining how their students compare, in terms of proficiency, with students at similar levels of instruction in other institutions.

VI. REFERENCES

The references listed below constitute a selected bibliography of materials in print on the topic of the ACTFL/ETS oral proficiency evaluation process. The numbers in brackets [ ] refer to references in *Resources in Education* (RIE), a bibliographic catalogue published monthly by the Educational Resource Information Center (ERIC). RIE catalogues and the *Cummulative Index to Journals in Education* (CIJE), also published by ERIC, are available in university libraries. RIE documents are contained on microfiche and generally found in university libraries. They also may be ordered from ERIC.

1. *ACTFL Provisional Proficiency Guidelines.* American Council on the Teaching of Foreign Languages: Hastings-on-Hudson, NY, 1982. [ED 230 072]

2. Buck, Kathryn, and David V. Hiple. "The Rationale for Defining and Measuring Foreign Language Proficiency Programs for Business." *Foreign Language Annals* 17 (1984): 525-29.

3. Canale, Michael. "Considerations in the Testing of Reading and Listening Proficiency." *Foreign Language Annals* 17 (1984): 349-60. See also James R. Child, Randall Jones, Judith E. Liskin-Gasparro, and Pardee Lowe, Jr.: pp. 361-89.

4. Ervin, Gerald. "Assessing Oral Proficiency: Some Practical Considerations." *Russian Language Journal* 34, No. 118 (1980): 7-19.

5. *ETS Oral Proficiency Testing Manual.* Educational Testing Service: Princeton, NJ, 1982.

6. Freed, Barbara. "Establishing Proficiency-Based Language Requirements." *Bulletin of the Association of Departments of Foreign Languages* 13, No. 2 (1981): 6-12.

7. Higgs, Theodore V., ed. *Teaching for Proficiency, the Organizing Principle*. ACTFL Foreign Language Education Series. National Textbook Company: Lincolnwood, IL, 1983.

8. _____, and Ray Clifford. "The Push Toward Communication," in Theodore V. Higgs, ed., *Curriculum, Competence, and the Foreign Language Teacher*. ACTFL Foreign Language Education Series. National Textbook Company: Lincolnwood, IL, 1982.

9. Johnson, Carl H., and Bobby W. LaBouve. "A Status Report on the Testing of Prospective Language Teachers for Initial State Certification." *Foreign Language Annals* 17 (1984): 461-74.

10. Kaplan, Isabelle M. "Oral Proficiency Testing and the Language Curriculum: Two Experiments in Curricular Design for Conversation Courses." *Foreign Language Annals* 17 (1984): 491-98.

11. Larson, Jerry W. "Testing Speaking Ability in the Classroom: the Semi-direct Alternative." *Foreign Language Annals* 17 (1984): 499-508.

12. Liskin-Gasparro, Judith E. "The ACTFL Proficiency Guidelines: Gateway to Testing and Curriculum." *Foreign Language Annals* 17 (1984): 475-90.

13. _____, and Protase E. Woodford. "Proficiency Testing in Foreign Language Classrooms," in Robert G. Mead, Jr., ed., *The Foreign Language Teacher: The Lifelong Learner*. The Northeast Conference on the Teaching of Foreign Languages: Middlebury, VT, 1982.

14. Lowe, Pardee, Jr. "Structure of the Oral Interview and Content Validity." [ED 223 108]

15. _____, and Judith E. Liskin-Gasparro. *Testing Oral Proficiency*. Eric Clearinghouse on Languages and Linguistics: Washington, DC, 1982. [ED 323 483]

16. Met, Myriam. "Listening Comprehension and the Young Second Language Learner." *Foreign Language Annals* 17 (1984): 519-24.

17. *POPT*. Publication of PROJECT OPT. Vol. I, No. 1 (March, 1984) and No. 2 (September, 1984). University of New Hampshire: Durham, NH.

18. Omaggio, Alice C. *Proficiency-Oriented Classroom Testing*. Language in Education: Theory and Practice, No. 52. Center for Applied Linguistics: Washington, DC, 1983. [ED 233 589]

19. _____. "Methodology in Transition: The New Focus on Proficiency." *The Modern Language Journal* 67 (1983): 330-40.

20. Shohamy, Elana. "Rater Reliability of the Oral Interview Speaking Test." *Foreign Language Annals* 16 (1983): 219-22.

21. Stevenson, Douglas K. "Foreign Language Testing: All of the Above," in Charles J. James, ed., *Practical Applications of Research in Foreign Language Teaching*. ACTFL Foreign Language Education Series. National Textbook Company: Lincolnwood, IL, 1983.

22. Wing, Barbara H., and Sandi Mayewski. *Oral Proficiency Testing in Russian*. ERIC. [ED 240 869]

23. Woytak, Lidia. "Reading Proficiency and a Psycholinguistic Approach to Second Language Reading." *Foreign Language Annals* 17 (1984): 509-18.

APPENDICES

1. The Interagency Language Roundtable Oral Proficiency Scale

S-0 NO PROFICIENCY

Unable to function in the spoken language. Oral production is limited to occasional isolated words. Has essentially no communicative ability.

S-0+ MEMORIZED PROFICIENCY

Able to satisfy immediate needs using rehearsed utterances. Shows little real autonomy of expression, flexibility, or spontaneity. Can ask questions or make statements with reasonable accuracy only with memorized utterances or formulae. Attempts at creating speech are usually unsuccessful.

Examples: The S-0+'s vocabulary is usually limited to areas of immediate survival needs. Most utterances are telegraphic; that is, functors (linking words, markers, and the like) are omitted, confused, or distorted. An S-0+ can usually differentiate most significant sounds when produced in isolation, but, when combined in words or groups of words, errors may be frequent. Even with repetition, communication is severely limited even with persons used to dealing with foreigners. Stress, intonation, tone, etc. are usually quite faulty.

S-1 ELEMENTARY PROFICIENCY (Base Level)

Able to satisfy minimum courtesy requirements and maintain very simple face-to-face conversations on familiar topics. A native speaker must often use slowed speech, repetition, paraphrase, or a combinaton of these to be understood by an S-1. Similarly, the native speaker must strain and employ real-world knowledge to understand even simple statements/questions from the S-1. An S-1 speaker has a functional, but limited proficiency. Misunderstandings are frequent, but the S-1 is able to ask for help and to verify comprehension of native speech in face-to-face interaction. The S-1 is unable to produce continuous discourse except with rehearsed material.

Examples: Structural accuracy is likely to be random or severely limited. Time concepts are vague. Vocabulary is inaccurate, and its range is very narrow. The S-1 often speaks with great difficulty. By repeating, such speakers can make themselves understood to native speakers who are in regular contact with foreigners but there is little precision in the information conveyed. Needs, experience, or training may vary greatly from individual to individual; for example, S-1s may have encountered quite different vocabulary areas. However, the S-1 can typically satisfy predictable, simple, personal, and accommodation needs; can generally meet courtesy, introduction, and identification requirements; exchange greetings; elicit and provide, for example, predictable and skeletal

biographical information. An S-1 might give information about business hours, explain routine procedures in a limited way, and state in a simple manner what actions will be taken. The S-1 is able to formulate some questions even in languages with complicated question constructions. Almost every utterance may be characterized by structural errors and errors in basic grammatical relations. Vocabulary is extremely limited and characteristically does not include modifiers. Pronunciation, stress, and intonation are generally poor, often heavily influenced by another language. Use of structure and vocabulary is highly imprecise.

S-1+ ELEMENTARY PROFICIENCY (Higher Level)

Can initiate and maintain predictable face-to-face conversations and satisfy limited social demands. The S-1+ may, however, have little understanding of the social conventions of conversation. The interlocutor is generally required to strain and employ real-world knowledge to understand even some simple speech. An S-1+ may hesitate and may have to change subjects due to lack of language resources. Range and control of the language are limited. Speech largely consists of a series of short, discrete utterances.

Examples: An S-1+ is able to satisfy most travel and accommodation needs and a limited range of social demands beyond exchanges of skeletal biographical information. Speaking ability may extend beyond immediate survival needs. Accuracy in basic grammatical relations is evident, although not consistent. May exhibit the commoner forms of verb tenses, for example, but may make frequent errors in formation and selection. While some structures are established, errors occur in more complex patterns. The S-1+ typically cannot sustain coherent structures in longer utterances or unfamiliar situations. Ability to describe and give precise information is limited. Person, space, and time references are often used incorrectly. Pronunciation is understandable to natives used to dealing with foreigners. Can combine most significant sounds with reasonable comprehensibility, but has difficulty in producing certin sounds in certain positions or in certain combinations. Speech will usually be labored. Frequently has to repeat utterances to be understood by the general public.

S-2 LIMITED WORKING PROFICIENCY (Base Level)

Able to satisfy routine social demands and limited work requirements. Can handle routine work-related interactions that are limited in scope. In more complex and sophisticated work-related tasks, language usage generally disturbs the native speaker. Can handle with confidence, but not with facility, most normal, high-frequency social conversational situations including extensive, but casual conversations about current events, as well as work, family, and autobiographical information. The S-2 can get the gist of most everyday conversations but has some difficulty understanding native speakers in situations that require specialized or sophisticated knowledge. The S-2's utterances are minimally cohesive. Linguistic structure is usually not very elaborate and not thoroughly controlled; errors are frequent. Vocabulary use is appropriate for high-frequency utterances, but unusual or imprecise elsewhere.

Examples: While these interactions will vary widely from individual to individual, an S-2 can typically ask and answer predictable questions in the workplace and give straightforward instructions to subordinates. Additionally, the S-2 can participate in personal and accommodation-type interactions with elaboration and facility; that is, can give and understand complicated, detailed, and extensive directions, and make non-routine changes in travel and accommodation arrangements. Simple structures and basic grammatical relations are typically controlled; however, there are areas of weakness. In the commonly taught languages, these may be simple markings such as plurals, articles, linking words, and negatives or more complex structures such as tense/aspect usage, case morphology, passive constructions, word order, and embedding.

S-2+ LIMITED WORKING PROFICIENCY (Higher Level)

Able to satisfy most work requirements with language usage that is often, but not always, acceptable and effective. An S-2+ shows considerable ability to communicate effectively on topics relating to particular interests and special fields of competence. Often shows a high degree of fluency and ease of speech, yet when under tension or pressure, the ability to use the language effectively may deteriorate. Comprehension of normal native speech is typically nearly complete. An S-2+ may miss cultural and local references and may require a native speaker to adjust to his/her limitations in some ways. Native speakers often perceive the S-2+'s speech to contain awkward or inaccurate phrasing of ideas, mistaken time, space, and person references, or to be in some way inappropriate, if not strictly incorrect.

Examples: Typically an S-2+ can participate in most social, formal, and informal interactions; but limitations either in range of contexts, types of tasks, or level of accuracy hinder effectiveness. The S-2+ may be ill at ease with the use of the language either in social interaction or in speaking at length in professional contexts. An S-2+ is generally strong in either structural precision or vocabulary, but not in both. Weakness or unevenness in one of the foregoing, or in pronunciation, occasionally results in miscommunication. Normally controls, but cannot always easily produce general vocabulary. Discourse is often incohesive.

S-3 GENERAL PROFESSIONAL PROFICIENCY (Base Level)

Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social, and professional topics. Nevertheless, an S-3's limitations generally restrict the professional contexts of language use to matters of shared knowledge and/or international convention. Discourse is cohesive. An S-3 uses the language acceptably, but with some noticeable imperfections; yet, errors virtually never interfere with understanding and rarely disturb the native speaker. An S-3 can effectively combine structure and vocabulary to convey his/her meaining accurately. An S-3 speaks readily and fills pauses suitably. In face-to-face conversation with natives speaking the standard dialect at a normal rate of speech, comprehension is quite complete. Although cultural references, proverbs, and the implications of nuances and idiom may not be fully understood, the S-3 can easily repair the conversation. Pronunciation may be obviously foreign. Individual sounds are accurate; but stress, intonation, and pitch control may be faulty.

Examples: Can typically discuss particular interests and special fields of competence with reasonable ease. Can use the language as part of normal professional duties such as answering objections, clarifying points, justifying decisions, understanding the essence of challenges, stating and definding policy, conducting meetings, delivering briefings, or other extended and elaborate informative monologues. Can reliably elicit information and informed opinion from native speakers. Structural inaccuracy is rarely the major cause of misunderstanding. Use of structural devices is flexible and elaborate. Without searching for words or phrases, an S-3 uses the language clearly and relatively naturally to elaborate concepts freely and make ideas easily understandable to native speakers. Errors occur in low-frequency and highly complex structures.

S-3+ GENERAL PROFESSIONAL PROFICIENCY (Higher Level)

Is often able to use the language to satisfy professional needs in a wide range of sophisticated and demanding tasks.

Examples: Despite obvious strengths, may exhibit some hesitancy, uncertainty, effort, or errors which limit the range of language-use tasks that can be reliably performed. Typically there is particular strength in fluency and one or more, but not all, of the following: has breadth of lexicon, including low- and medium-frequency items, especially

sociolinguistic/cultural references and nuances of close synonyms; employs structural precision, with sophisticated features that are readily, accurately, and appropriately controlled (such as complex modification and embedding in Indo-European languages); has discourse competence in a wide range of contexts and tasks, often matching a native speaker's strategic and organizational abilities and expectations. Occasional patterned errors occur in low-frequency and highly complex structures.

S-4 ADVANCED PROFESSIONAL PROFICIENCY (Base Level)

Able to use the language fluently and accurately on all levels normally pertinent to professional needs. An S-4's language usage and ability to function are fully successful. Organizes discourse well, employing functional rhetorical speech devices, native cultural references, and understanding. Language ability only rarely hinders him/her in performing any task requiring language; yet, an S-4 would seldom be perceived as a native. Speaks effortlessly and smoothly and is able to use the language with a high degree of effectiveness, reliability, and precision for all representational purposes within the range of personal and professional experience and scope of responsibilities. Can serve as an informal interpreter in a range of unpredictable circumstances. Can perform extensive, sophisticated language tasks, encompassing most matters of interest to well-educated native speakers, including tasks which do not bear directly on a professional specialty.

Examples: Can discuss in detail concepts which are fundamentally different from those of the target culture and make those concepts clear and accessible to the native speaker. Similarly, an S-4 can understand the details and ramifications of concepts that are culturally or conceptually different from his/her own. Can set the tone of interpersonal official, semi-official, and non-professional verbal exchanges with a representative range of native speakers (in a range of varied audiences, purposes, tasks, and settings). Can play an effective role among native speakers in such contexts as conferences, lectures, and debates on matters of disagreement. Can advocate a position at length, both formally and in chance encounters, using sophisticated verbal strategies. Can understand and reliably produce shifts of both subject matter and tone. Can understand native speakers of the standard and other major dialects in essentially any face-to-face interaction.

S-4+ ADVANCED PROFESSIONAL PROFICIENCY (Higher Level)

Speaking proficiency is regularly superior in all respects, usually equivalent to that of a well-educated, highly articulate native speaker. Language ability does not impede the performance of any language-use task. However, an S-4+ would not necessarily be perceived as culturally native.

Examples: An S-4+ organizes discourse well, employing functional rhetorical speech devices, native cultural references and understanding. Effectively applies a native speaker's social and circumstantial knowledge. However, cannot sustain that performance under all circumstances. While an S-4+ has a wide range and control of structure, an occasional non-native slip may occur. An S-4+ has a sophisticated control of vocabulary and phrasing that is rarely imprecise, yet there are occasional weaknesses in idioms, colloquialisms, pronunciation, cultural reference or there may be an occasional failure to interact in a totally native manner.

S-5 FUNCTIONALLY NATIVE PROFICIENCY

Speaking proficiency is functionally equivalent to that of a highly articulate, well-educated native speaker and reflects the cultural standards of the country where the language is natively spoken. An S-5 uses the language with complete flexibility and intuition, so that speech on all levels is fully accepted by well-educated native speakers in all of its features, including breadth of vocabulary and idiom, colloquialisms, and pertinent cultural references. Pronunciation is typically consistent with that of well-educated native speakers of a non-stigmatized dialect.

2. The ACTFL/ETS Oral Proficiency Scale

Novice—Low

Unable to function in the spoken language. Oral production is limited to occasional isolated words. Essentially no communicative ability.

Novice—Mid

Able to operate only in a very limited capacity within very predictable areas of need. Vocabulary limited to that necessary to express simple elementary needs and basic courtesy formulae. Syntax is fragmented, inflections and word endings frequently omitted, confused, or distorted, and the majority of utterances consist of isolated words or short formulae. Utterances rarely consist of more than two or three words and are marked by frequent long pauses and repetition of an interlocutor's words. Pronunciation is frequently unintelligible and is strongly influenced by first language. Can be understood only with difficulty, even by persons such as teachers who are used to speaking with non-native speakers or in interactions where the context strongly supports the utterance.

Novice—High

Able to satisfy immediate needs using learned utterances. Can ask questions or make statements with reasonable accuracy only where this involves short memorized utterances or formulae. There is no real autonomy of expression, although there may be some emerging signs of spontaneity and flexibility. There is a slight increase in utterance length but frequent long pauses and repetition of interlocutor's words still occur. Most utterances are telegraphic and word endings are often omitted, confused, or distorted. Vocabulary is limited to areas of immediate survival needs. Can differentiate most phonemes when produced in isolation but when they are combined in words or groups of words, errors are frequent and, even with repetition, may severely inhibit communication even with persons used to dealing with such learners. Little development in stress and intonation is evident.

Intermediate—Low

Able to satisfy basic survival needs and minimum courtesy requirements. In areas of immediate need or on very familiar topics, can ask and answer simple questions, initiate and respond to simple statements, and maintain very simple face-to-face conversations. When asked to do so, is able to formulate some questions with limited constructions and much inaccuracy. Almost every utterance contains fractured syntax and other grammatical errors. Vocabulary inadequate to express anything but the most elementary needs. Strong interference from native language occurs in articulation, stress, and intonation. Misunderstandings frequently arise from limited vocabulary and grammar and erroneous phonology but, with repetition, can generally be understood by native speakers in regular contact with foreigners attempting to speak their language. Little precision in information conveyed owing to tentative state of grammatical development and little or no use of modifiers.

Intermediate—Mid

Able to satisfy some survival need and some limited social demands. Is able to formulate some questions when asked to do so. Vocabulary permits discussion of topics beyond basic survival needs such as personal history and leisure time activities. Some evidence of grammatical accuracy in basic constructions, for example, subject-verb agreement, noun-adjective agreement, some notion of inflection.

Intermediate—High

Able to satisfy most survival needs and limited social demands. Shows some spontaneity in language production but fluency is very uneven. Can initiate and sustain a general conversation but has little understanding of the social conventions of conversation. Developing flexibility in a range of circumstances beyond immediate survival needs. Limited vocabulary range necessitates much hesitation and circumlocuation. The commoner tense forms occur but errors are frequent in formation and selection. Can use most question forms. While some word order is established, errors still occur in more complex patterns. Cannot sustain coherent structures in longer utterances or unfamiliar situations. Ability to describe and give precise information is limited. Aware of basic cohesive features such as pronouns and verb inflections, but many are unreliable, especially if less immediate in reference. Extended discourse is largely a series of short, discrete utterances. Articulation is comprehensible to native speakers used to dealing with foreigners, and can combine most phonemes with reasonable comprehensibility, but still has difficulty in producing certain sounds, in certain positions, or in certain combinations, and speech will usually be labored. Still has to repeat utterances frequently to be understood by the general public. Able to produce some narration in either past or future.

Advanced

Able to satisfy routine social demands and limited work requirements. Can handle with confidence but not with facility most social situations including introductions and casual conversations about current events, as well as work, family, and autobiographical information; can handle limited work requirements, needing help in handling any complications or difficulties. Has a speaking vocabulary sufficient to respond simply with some circumlocutions; accent, though often quite faulty, is intelligible; can usually handle elementary constructions quite accurately but does not have thorough or confident control of the grammar.

Advanced Plus

Able to satisfy most work requirements and show some ability to communicate on concrete topics relating to particular interests and special fields of competence. Generally strong in either grammar or vocabulary, but not in both. Weaknesses or unevenness in one of the foregoing or in pronunciation result in occasional miscommunication. Areas of weakness range from simple constructions such as plurals, articles, prepositions, and negatives, to more complex structures such as tense usage, passive constructions, word order, and relative clauses. Normally controls general vocabulary with some groping for everyday vocabulary and relative clauses. Normally controls general vocabulary with some groping for everyday vocabulary still evident. Ofen shows remarkable fluency and ease of speech, but under tension or pressure language may break down.

Superior

Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social, and professional topics. Can discuss particular interests and special fields of competence with reasonable ease. Vocabulary is broad enough that speaker rarely has to grope for a word; accent may be obviously foreign; control of grammar good; errors virtually never interfere with understanding and rarely distrub the native speaker.